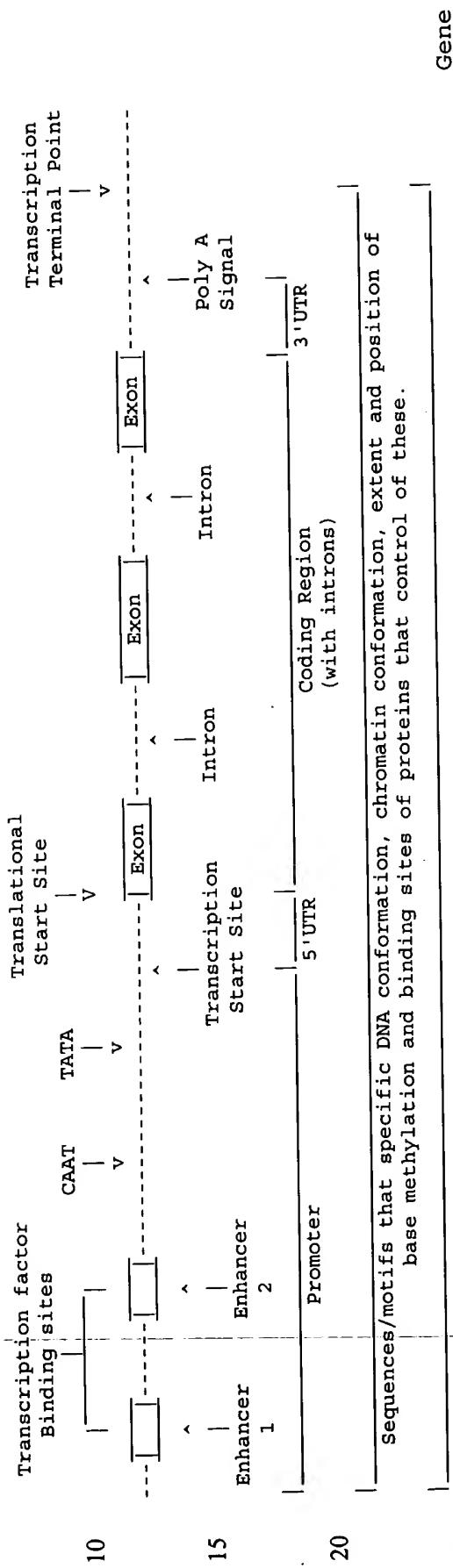


09/724857

FIGURE 1  
SCHEMATIC OF A GENE

5



25

8

**FIGURE 2**

5	-4242	GCATGCACTG	CCACAAGTAG	TGAACTCATG	GTTTTACCTC	CTCAAGTAGA
	-4192	AAACCTTTG	AGTGAATTG	AAGATTTATT	CTCCCAAGAA	GGACCCATTG
	-4142	GGCTTCCTCC	TCTTAGGGGG	ATAGAACATC	AAATTGACTT	TATACCGGGG
	-4092	GCAAGCCTAC	CAAATAGGCC	TCCTTATAGA	ACCAACCCCG	AGGAAACAAA
	-4042	GGAGATAGAA	TCACAAGTTC	AAGACTTGTT	GGAGAAGGGT	TGGGTTCAAA
	-3992	AGAGCCTAAG	CCCTTGTGCT	GTACCTGTCT	TGTTGGTGCC	AAAAAAAGAT
	-3942	GGAAAATGGC	GTATGTGTTG	TGATTGTTAGA	GCAATCAACA	ACATCACCAT
10	-3892	CAAGTATAGG	CATCCAATCC	CAAGGCTTGA	CGATATGCTT	GATGAATTGC
	-3842	ATGGGTCAAC	TCTATTCTCC	AAAATGACC	TTAAAAGTGG	ATATCACCAA
	-3792	ATTCGAATCA	AGGAGGGTGA	TGAGTGGAAA	ACCGCTTTA	AGACCAAATT
	-3742	TGGATTATAT	GAGTGGTTGG	TGATGCCCTT	TGGTCTTACT	AACGCTCCAA
	-3692	GTACATTCTAT	GAGGCTTATG	AATCACACCT	TGAGGGATTG	TATAGGTAAA
	-3642	TATGTAGTAG	TTTATTGTA	TGATATCTTA	GTATATAGTA	AAACCCTAGA
15	-3592	AGACCATCTA	AGTCACCTTA	GGGAAGTTCT	TCTAGTTCTT	AGGAAAATA
	-3542	GTCTTTTGC	CAATAGGGAT	AAAGTGTACCT	TTTGTGTTAGA	TAGCGTAGTC
	-3492	TTTTTAGGCT	TTATAGTAAA	CCAAAAGGGG	GTGCATGTAG	ATCCCAGAGA
	-3442	AATCAAAGCC	ATCCGCGAGT	GGCCAACCTCC	ACAAAATGTA	AGTGATGTGA
	-3392	GAAGTTTCA	TGGGTTAGCT	AGCTTCTATA	GAAGGTTGT	TCCCAATT
20	-3342	TCTAGCCTAG	CTTCTCCCTT	GAATGAACCTT	GTAAAAAAAG	ATGTTGCATT
	-3292	TTGTTGGAAT	GAAAAGCATG	AGCAAGCCTT	TCAAAGGCTA	AAAGCTCACT
	-3242	CACCAATGCA	CCCATCCTAT	CTCTTCCAAA	TTTTTCCAAA	CTTTTGGAGA
	-3192	TAGAGTGTGA	TGCATCGGGA	GTAGGCATAG	TGCGGTTTG	TTGCAAGGTG
	-3142	GACACCCCTT	GCTTATTTA	GTGAAAAAACT	CCATGGTGCC	ACCCCTCACTA
25	-3092	CCCCACCTAT	GACAAAGACT	CTATGCTCTT	GTGCGACCCCT	AAAGACTTGG
	-3042	GGAACACTAC	CTTGTGCCCC	AAAGAATTG	GnTATCCATA	GTGATCACGA
	-2992	GTCTTTAAAA	TATTTAAAGG	GCCAACACAA	GCTCAATAAG	AGACATGCTA
	-2942	AATGGATGGA	ATTCTTGAA	CAATTTCTT	ATGTCATCAA	ATACAAGAAA
	-2892	GGGAGCACCA	ATATAGTGGC	CGATGCTCTT	TCTAGACGGC	ACACTCTCTT
30	-2842	TTCAAAACTA	GGTGCCCCAA	TTCTTGGATT	TGACCACATA	AGAGAGCTTT
	-2792	ATCAAGAAGA	TCAAGAACCTC	TCATCCATCT	ATGCCAATG	TCTACATAGA
	-2742	GCACAAGGAG	GTTACTATGT	GTCCGAGGG	TATCTTTTAA	AAGAAGGAAA
	-2692	ACTTTGCATT	CCCCAAGGAA	CACATAGAA	ACTCCTGTC	AAAGAATCAC
	-2642	ATGAAGGGGG	ACTCATGGC	CATTITGGAG	TTGATAAAAC	TCTAGACTTT
35	-2592	TAAGGCAAA	ATTTGTTGG	CCACATGA	GGAAAGATGT	CCACGACATT
	-2542	GTCTAGAGTA	TCTCATGTTT	AAAAGCAAAG	TCTAGAACAA	TGCCGCTGGA
	-2492	CTCTACACCC	CTTGTGCGAT	TGCAAAGCTC	CTTGTGAAGA	CATTAGCATG
	-2442	GATTTCATTT	TAGGACTTCC	TAGGACTGCA	AGAGGCCATG	ACTCTATCTT
	-2392	TGTGGTAGTG	GACCGTTTTA	GCAAAATGTC	TCACTTTATT	CCATGCCACA
40	-2342	AACTAGATGA	TGCTCAAAAT	ATTCTAAAC	TCTCTTTAG	AGAAGTGGT
	-2292	AGACTCCATG	GTCTCCCTAG	AAAGTATAGT	TCCGATAGAG	ATCACCTTAA
	-2242	ATATATAATT	ATACACTTGT	TTTTTTCTC	TTTTTTATT	TATCAAGTAA
	-2192	AAAGTATTG	TTCTAGATT	TTATGAGTAT	ATACTTACTT	TCTGTATTTC
	-2142	ATTTCTTTCT	ATTTTTATG	ACGATGAAAT	TTCTTATT	ATCCAGACTT
45	-2092	TTCATATATA	TTTTTATTTC	TTTCCATCT	AGATGCTCTG	TACTTTCTT
	-2042	CAGTTGAAAT	TTCCACTCTC	CAACAAACAA	TCATTCAAGT	TTTGATAAC
	-1992	ACTGTGACGT	TAACCAAGTTA	AAATAAGAAA	ATCATGTAAT	ATAAATTATT
	-1942	TCAGTAGATA	TTTTAGAATT	ACAAATACGA	AAATAATT	ATTTAAAAAA
	-1892	ATTATTAAAC	AATGAATT	TTTGGAAATT	AATATAAAAC	TTAGACTTGT
50	-1842	GGTTTCTTC	TTCAGTCAAA	ACCTTTCT	ATTGTGTTGGC	GTGTGCGTGA
	-1792	ACATCGAATT	TGGGTGCTTT	ATGCCGCTT	ATCTTCATCT	GCACCTTCAA
	-1742	ATTAATAATT	TAATTCCGGA	AAATAATAAA	CCCACACACT	GTTTTATGCA
	-1692	TATATTAAAGA	AAATAAAAG	AGAACTATT	AAAGAATAT	AAAATAATAA
	-1642	ATGTAACAA	TGATGTCACT	AAAGAAGAAA	AAAATTAACA	AGAATTGTAA
55	-1592	TATATTCTT	TATGAAATGT	TTTGTGCTT	ACCGAGAGAG	GTGCAACATG
	-1542	ATACACGCAA	GCATCTAACT	AGTTTGGTAA	TCCTTTCA	ACATCGnTAA
	-1492	GCACATCACA	CTAAAATTAC	TTTAAATAGA	AAATTAGAT	TCAATTGGAT
	-1442	GACATTAATT	TATAATACTC	TATCCAAAT	TATAACTATA	AATAAAAAGT
	-1392	TATTTTTAGA	AAATAAGTAA	TGAAAATTAA	ATTCTAAAT	TTATAACACT
60	-1342	TTTATGCTGT	GTGTTGTTCG	AAGCATAGAA	AAATAAAAAG	TTATTGTTGG

-1292 GAATGAAAAG TGAAGAAAAT CATGTAATAA AAACAAAATG ACACGACAAT  
 -1242 CAAAAAAAAGA GTTTTCATGC AAAACTTTT TCAAAATTAA CACTTTATG  
 -1192 ATGTGTTTGT TTCAAGTGT AGAAAACGA AAAGTTATTA TTGTTAATGA  
 -1142 AAAGCGAAGA AAATCACGTA ATAAAAACAA AGCAAGATGG CACGACAATC  
 5 -1092 AAAAAAAAGT TTCTACACAA AACTTTATTC AAAATTACA ACACTTTAT  
 -1042 GTTGTGTTT GTTCCGAGG TATAGAAAAA CAAAGAATTA GTGTTGGTAA  
 -992 TGAAAAGTGA AGAAAACCAT GTAATGAAAA CAAATGGCA CGACAATCAA  
 -942 AAAAAAGTTT CACGCAAAAT TTTCTTCAAA ATTTATAACA TTTTCATGTT  
 -892 GTGTTGTTT CAAAGCCTAG AAAAACGAAG AGTTACTATT GGTAATGAAA  
 10 -842 AGCGAAGAAA ACCACATAAT AAAAACAAAA TGGCAGGACA ATCAAGAAAA  
 -792 AGTTTTCA CAAAACTTTT TTCAAAATT ACTATGTTA TTTCGAAATT  
 -742 TAGAAAAACG AAGAGTTATT ATTAGTAATG AAAAGCGAAG AAAACTACGT  
 -692 AATAAAAAC AAAATGGCAC GACAATAAA AAAGTTTCA CGCAAATTT  
 15 -642 TCTTGGTGCG CAGAAAAGTTA TATATATTAA TTAATTAAATT TTCAATTACT  
 -592 TTTTCCCTT TTTATTAAAGTTAAATTAA TTATTATTTT CATTAAAAT  
 -542 ATAAATATTA TTTAAATATA AAAATATAAA CCTTAATCAA ACAAAAGCCT  
 -492 TAATCTAAA TTTACAACAC TTTAACCTT AAAATTAAC TTAAAAGGAA  
 -442 AATGATAGTG TGACAACAA AAAAGTTGTA TACAACCTG TCATAGGTTT  
 -392 AGAAAATAAT ATATATAATA AAGAGTAAAT TTGTAATTAA ATGATATAAA  
 20 -342 AAAGTATTAA AATAATAATA TTTAGAGTAG TAATATGGTT GTATAAAAAA  
 -292 ATGTGGTTGT CCATATATCA TTATTCACTT TAAATATCA TGACAAATAT  
 -242 TTTCACCGAA AGATGGAAAG AACGAAAAGA GCGTTGGATA ATGGAAAAAT  
 -192 ACAAGCAATC TCCCTCCAGT ACTTGCATA ACATTTGTA TTAGTGTGAA  
 -142 GTTTTATC ATATATATTT AGAATATAGG AAAATTTAG AATCACGTGG  
 25 -92 ATAGCTATAT AATAGTAATA TTTAAATTAA TAATGTAGTT GATTTATTT  
 -42 GTCAACTGGT ATACATAAAAT ATGTGTTGAT AGTGGGTGAC TTGTGGCTTA  
 9 AAGAAATGTC CAGAGGCTGA CAACAACCT GCACAGACTA GCGTAAAC  
 57 **ATG AAG TCC AAT TTT GCT ATT TTC GTA GTC TTT TCT CTT CTT CTT**  
 1 M K S N F A I F V V F S L L L  
 30 102 CTG GTACCTCTTCATCTCTACAAAAACTCTGTTGCTCTTCACCTCTGTTGTA  
 16 L  
 160 ATTTGTTTACACTTTGAAAATTGAAGCTGATATATATGTAACAAACCTTCAGTTT  
 219 GTCTGCACTGAAACTGATAGAAAATATACGTTTGTTGATATATAG GTT GGC  
 17 V G  
 35 274 AGT TGC AGC TGC GCA AGA AAA GAC ATG AGA GGG TAT TGG AAG GAT  
 19 S C S C A R K D M R G Y W K D  
 319 ATG ATG AAG GAG CAA CCT ATG CCA GAA GCA ATC AAA GAC CTT ATT  
 34 M M K E Q P M P E A I K D L I  
 GAG GAT TCA GAA GAA GTG TCA GAA GCA GGG AAG GGT CGT TTT GTT  
 40 49 E D S E E V S E A G K G R F V  
 AGG GAC TTC GAT GTA AAG CCT AAT GTC ATA TTA TAT CAC ACA CAT  
 64 R D F D V K P N V I L Y H T H  
 GTT GTG CCC ATG AAG CAG AGG CAG AAG AAT AAA GAT TGA  
 79 V V P M K Q R Q K N K D •  
 45  
 493 AGACTATGTGATTGGCAGTTTCAGACTTATTTGGCACCAAATTATGATGCTCTGTTGCTG  
 555 TTTCAAAATTGTACTCAAACCTTGACCCCTTGACAGCATCTGCTCTTTGGTCTTGCT  
 617 GAATTTGTCACAGTTACTGTACGAATAGTTCTCTTCATAATAAGCAACTTTCCCTCT  
 679 C

**FIGURE 3**

101001 CAAAACAAAAGCAAATGCCGGTTTCTTATTATTATTCGAACTTAGAC  
5 100151 CTTTTGTAACGTTCTTAATTTTTCTTGTAAAGAACCCATTAT  
100201 ATCTTAGCTAAATATTACCTCATTTGTTATGAGCTAACCCACCCAA  
100251 AAATATTGAGTTGCTTCGGATTTAAGCCAAGCAAGTATTAGAT  
100301 ATATTAAGGAAAATGAATGAAAGGACAAAAAAATATAACGACAATATT  
100351 TGAATACTGATATTATCTCCATTCTCAAATATTTTGATTATTGTGAC  
10 100401 AATATTTGGTGTTCACATCTTGAGGACATGAAATGATA  
100451 ACATATATATGAACGAGTATAATACATTCTGTTCAATTCAAATAAT  
100501 GTCAATTATGCTAACATTTTATTAAAGGATCCTTATAAGATTTC  
100551 AGTGTTATTACCATGGTACTGAAAGTCGGATGCTATATATATA  
100601 TATATATATATATCAAATGACACTGAAGAATTGAACTAAAA  
15 100651 CTAAAAACGTAAAATAAGAATTTCAAAATCAAATTTATATA  
100701 AAAATATAGATAAAATGTTAATATAGTACAACCTCTATTCAAACAGAG  
100751 AATAAAATCTCTATAGACAGTGAATATCCATTATAACGAGCAATAGT  
100801 TGTAATGTTGCACTGACAAAAGAGAATTGTAATATTGTGATGATTGAG  
100851 AAATCTAAGTGACTTGAATTAAAAGGCTAATTCCAACAAGTACATGTA  
20 100901 GAAGTTGACTATAGCTATATATTACTACAAATTGATCATTCAGAAAG  
100951 ACATTTAAATTAAGATATGCATGCATGACTTGATTGAACCCACTCGCTT  
101001 GCTTCGTGCCATTGACAAGATGTTACTTTAAATGCAAGTAAATTATG  
101051 GATATACTCTCTGTATTGTTAGTAAATTTCAAGAAATTGTTT  
101101 TTTTCAAAATCAAATGATATTAAATTAAATTCAATATAGAATTAA  
25 101151 ATTAAATTAATTGAAAGATTATGCTGCAGATTAGATTACCATG  
101201 GTGAAATCATGTTAGGTTAAATAAAATGATGTTGAGTTAGGAAAAAA  
101251 AAAAATCTTAAATCTTATGTAAGAAATGTTAAACTTCAATTATAAAAAA  
101301 TATGAAGCAGTATTATAAGATGTTAACTAATGAAATAATTGTTT  
101351 GGATGAAATTCTTGCAATGTTCTAAAGGAAATTGTTAAATATGAAAAAATT  
30 101401 AACATTCAATTGTTAGGTTAAAGAAATATGAGTTGTTAGATAA  
101451 ATAATACCTAAATTAAAGAATTGTTAAAGTTACTGCACCTCAAATATG  
101501 TTATTTTCTTTTATTAAATCAGCAACATTCTAAATGATTAT  
101551 TTTCTTAAAAATTGAAAGAAATTGCAAAATATGTTAAATTAA  
101601 ACGAATTAAAGAAAAACTTGAAAGATATGATATGCTTATAAAAAAA  
35 101651 ACTTGGTGGCGTACCTACTAAATATGATCACATTAGAGATTGTATCCTT  
101701 TAGCATATAGTATGTTAGTATGATATCTATATTATTATTAAAGAGC  
101751 ATATTCAATATAGTATTATGTTAATTACAATAACGTTCAATTG  
101801 TTATGTTAGTTTAGAAAATTGCGTGTGCATATCAATGTGAGAAA  
101851 GCGACTCCACATGTGAGATGTTGGCTGAGAAAGCTTCTGCACTTGGTC  
40 101901 GGAACTACTCATGGACTAGAATGCAATCCATCTATTCAAAGAAAAGCAG  
101951 TTGTCCATGCATGCCTCGGTTTCACATTGGAAGCAGCGCAACAATGT  
102001 CTTACATAATATGCATCGATCACTCTGCAACCAATTCAAGTACATAG  
102051 ACCATGACATCAAAACATTATCACACCGAGAAGAAAGAACGTCAATT  
102101 GTAACTTAAATGGCGTTATGCCTCGGTGAATTCTCTAAAGAGTTCTCC  
45 102151 AAATTAAATGATTCTCTGTTTAACCTTTCGCCAAAGAATCATACT  
102201 ATAGATTGACACCATTCAACTTACAAATACAAGTGAATAAAATAATT  
102251 CAAGCTGAAAGGAATTAAATCATGATCTAAACCTAAACGACAATTCTT  
102301 CACAAGTGAGAATCACTAAATTGACTACCCCTGGTCGATATACATCATT  
102351 GTTGTAAATCTGAAAATTGGTTGGATTGATCTGATATGTCATTCA  
50 102401 AAAACTTGTATTATTTAGAATTTCACCTGCGCAACAGATAAATCATC  
102451 ATCTATTAGAAAATTTCATTGCAACCAATTAAATCAGGGAAAAGGT  
102501 GAAATCACATATCTTACACTCTTATTAAATTAAACGCCATAATATA  
102551 ACAAAATTCTAAATACCACTTATGAGAAGCACTAAGATCACCTTTCTT  
102601 TATGACTTTCTTCTAAAGCTAAGCTGGTAGTCATGACTCATGATTATCC  
55 102651 TTTCTTAATGGAAATTGTGGAAGCGGTTCAATCTTAGACAAAAT  
102701 TCCATGGCCACTAAAAGTTAGCAAAGTTAAAGTTAAAAAATATG  
102751 AGTGACTTGGCCATATGCCATTGTTGAGATCATAACAAGAGAAATAA  
102801 TAGTTTATTGAAGTTAGATCATAATCACAAATACATCATTGCTTCA  
102851 ACATTTCCATGGATTGAGAGGATCAACTTCAACTAATGGGGGTG

102901 TTATTCATCCATTGCTCTAGCCAATTAAGCAGTTAGGTATTTGTGA  
 102951 CTCTAGTAGTTGCCAAATCAATCTTAATATTCAACATGTTGAATTCTA  
 103001 ATTACGTATAGATAAAATGACTAGATAACACGTGGCTTGGTTTATCAGG  
 103051 AAAGTTTCCAATCATATATGAATGTAGAATAGTGTCTTCATTAAT  
 5 103101 TATTAATTAGCATCTCACCATCTGAGACTGGGAGCATGTGACAAGTTGAC  
 103151 ATGTGTATTAAGAGAACTTTGAGAAAAACACTTTATGATACTCCCACATCT  
 103201 GAGACTGGGATGAGTACCATTTATAAAAATATGAGTAGTGAAAAAAATAT  
 103251 TCAAAAAAAATTCTAACATGTCTTAAACATTAAACCTTATAATT  
 103301 AACAAACATTTCCAATATGCGTTATGAAAACCTTTATAAAACTTTTAT  
 10 103351 AACATGCTTTGAAAATTTTATAAAATCTGTTAGAAACAAAGTGT  
 103401 ACTTTGAAAATAGACAAATGAAGTGTATTAGTATATCATA  
 103451 AGTCTTAACGTGGTTGTTGAATTTTATTTATATACTTGTCAAATAA  
 103501 AACTAAATAAAATTAAATTATTTATAATCATGAAGATAATATTAC  
 103551 ATAAAAGATAAAATATAAAATCAACAAATTATTTATAATTGTTAATAAAATAC  
 15 103601 TTTGAGCTCTTCTCATAGACTTTCCAGCTCCATCTAGAAAATCACA  
 103651 TAAATTAAAAGATAAAATAACCGAATAAACATAGTTCACATTCTAACCTT  
 103701 AGTCTTAGATTGTTAATTTCAAAAGTTAGGTATTGTATATGTTT  
 103751 TTTTATTGGGTTGCTAGATTTGATCCAAGAAGAAATGACGGGTTGTAGT  
 103801 ATAGATGGTTGTTGAGTTTTCCCCTGGTTACTCGTTGGTT  
 20 103851 TGTCCCCAGAATTGTTCTTGTACTCGCTGGTTATGTCCTACAAAGTCC  
 103901 ACGACCATTGCCGGCTCTTGTATTCACCTGAATTCTAAATTGATTG  
 103951 ATGAAAAAAAATGTATCTCTAAAGCTTACATTAGTACCAAAAATAACTAT  
 104001 ATCATTACTACATAAAATAGTCTTGGTTTCCAAAGTATTGTTGATA  
 104051 TATGTTAAGAGTTCGAAATAGACACATAGATAATGTTGAAATGGGACC  
 25 104101 TCTCACATAATTATCTCTTCTCTCATTTCTCTACCTCTCAAGTTTC  
 104151 CAATCCCACCCCTAAGGTAAATTATTTCTTAACCTAACGTAATTGTTAAC  
 104201 AAATCTTAACTAGCTACAAATGTGTATTACAAGTCTTAAATAAAACCTA  
 104251 CTTTAATTCAAAGGTATAAACCTCTTAAATTGATACTTACTTAGTATC  
 104301 GATCGGTCTAGTTAGGGTTGGACACACACCATCATGGGACGAAAATT  
 104351 AGTCATTCTACGGTGTCAAGACACAAATCTGGACTCGATGTGGATATG  
 104401 ACACCTTACATAACTTTAACCTACAAAAACTAACTATTAGGAGGAAG  
 104451 AATCGGAATCTGCATATCAACATACAATAGACTATAGTAACTTAGATT  
 104501 GATCTAACATGGCTCTCAACTAACATAAGTAGGCCACTACCAATAAT  
 104551 GAAATCATAAGACATTATTAAATTCAATGTTCTAAAATACTTTGGT  
 30 104601 TATGTGTCCCGTAGAGCTAATGTGCACACACAATGAAAGTTGACCGTT  
 104651 CACTTGTCCACTTTATGATCTTCTTTAGGTTAAATCCAACCTT  
 104701 TAATCTCATTTGTATCAAACAAACTTTGGCTGTCTTTCTATAAT  
 104751 TTAAAGTAACCTCACGGAGAAAAGCCAACATTCTTCTTGTCTTATTC  
 104801 TTTTTAAGAAAATGAATTCAAGGGGACCCCAATTAAAAGGAAAACCA  
 40 104851 AAACCTTCTTCTATGTATTACTTGAAGTTCTATGTAATCAACAA  
 104901 TCCTAACAGTAGAGAATAAAAACATCGTTGGGAGGTTTATTTAGC  
 104951 ATATGAGAAATGTTCTAAAATTGTTTACACAAAATTAGATTCTTT  
 105001 CCTCTGTCAATGGAGCTATACACTTGTCAATTGCTTAACCCCTTGCGG  
 105051 GAAGATTGTATGAAACAGTTAATGGAATTCTAGTTGCAATGTCAAG  
 45 105101 TTAAATATGTTGTCCTATACCTTATTGAACTTATAATCTTGTAT  
 105151 AGAATTATCTACTTTAGTATTTACATTAACATAATCTATAGAATTCTT  
 105201 CTTTGTCTATACAAATTAAACAGTAATATATTCTTAATACATATTAAA  
 105251 ATGGTGGTGTGCTATCTGAGCTGAATAGTTGATTGCTCAGAGAAGAA  
 105301 TAGACAAAATCTTACTTAAGAGGCCACCACTCTGAAAATTAGACAA  
 50 105351 GAAAATTAACAAAATTAGGTACACATATTATCATTATATATGCA  
 105401 CAACACAAAGTTGACCTGCAATGTACTATTGAATAAAATAAAATGC  
 105451 AAGAAGAGAGGGATTACTGTACCAAGAAAACAACCTCCTCTAAAC  
 105501 AGGTCTCTATATATATAACCTTACACCTAAAGAATTACACAGATCAA  
 105551 GAAAATCCTCAAAACAAAAGTTAAAGCAGAC **ATG** AAG CAA CAG CAA  
 55 1 M K Q Q Q  
 105599 CGT TAC TTG GTC GTC TTC ATC GTC CTT TTA AGC TTT CTT  
 6 R Y L V V F I V L L S F L  
 105638 CTG GTAAAGCTTCTCCTTAATTATATTAAACCTAATTAAAGATCTCATATA  
 19 L  
 60 105691 TCTGAATGTTGTATATATTGTTGGTATAG TTT GTG AAT CTG AGT  
 20 F V N L S

105736	GAA	GGA	AGA	ACA	GGA	GGA	GTT	GCA	GAA	GAA	TAT	TGG	AAG
25	E	G	R	T	G	G	V	A	E	E	Y	W	K
105775	AAG	ATG	ATG	AAG	AAT	GAA	CCG	TTG	CCT	GAA	CCA	ATC	AAA
38	K	M	M	K	N	E	P	L	P	E	P	I	K
5	105814	GAG	CTT	CTC	AAC	AAT	CCT	TTT	AGG	ACC	GCA	CAA	GAG
	51	E	L	L	N	N	P	F	R	T	A	Q	E
105853	TTC	ATC	CAG	AAT	TTC	GAC	ACC	AAA	TCT	GTT	GTC	ATC	ATC
64	F	I	Q	N	F	D	T	K	S	V	V	I	I
10	105892	TAC	CAC	AAT	CCT	AAT	GAA	TAATGAA	TCAATGAGTCTCTCATATAG				
	77	Y	H	N	P	N	E	•					
105934	ATATCTATGACTTTA	ATTTGTTATG	TATGGATCG	ACTTATACGTGCA									
105984	CGTATATGTTATTA	ATTAAGAAAAGAAAAG	CTGCTTGAGTTGTTGTT										
106034	ATACACGTATA	ACTAAATATGTTCTGTT	TAGTG	CAGAAATGTTAACCC	CTAG								
106094	CTATAAGGGAT	TTTTGTTCTTTTTT	TTTTGTT	ACCATTAATG	TGAGTGA								
15	106144	GTGAGTTTGTTGATGAA	AAATTAGATTTG	CTTCACATTTGTTT	GATA								
	106194	TATATAAATCAATATA	ACTGTGCC	TTCTG	TGCTT	GTTCTT	ATATTATTT						
106244	TGTGACATTA	ATTAATTATCTT	ATC	AAAATT	TTTATT	TATTAA	ACTGTGTC						
106294	CTATGGAAAAGATG	AAACAATATGAG	TTAAC	CTC	ATCT	CAAGGAGATT	C						
106344	TTTTGTTTGT	TTTGT	TTT	TTT	TTT	TTT	TTT						

20

**FIGURE 4**

1 AAGCTTACAAATGCCCCAAAGATGAAACCACGTTATTATTAGTAAATCCTGAAAAGG  
 61 TTAACGCTTCTGTCCTGAAATTCTAAACCATCTGAAATATCTAGTGGTTAAAATGGAG  
 5 121 ACTTGAGGATATAGTCTCTGAAACCAGCTGCACGGCTGAGTTAGATAACATTACTGAAT  
 181 TTCTACGGGAGCGGTGAAATCACTTCCGGCTTAAAGAAGAAGCCTACACCGGGCACC  
 241 TTCTTACGCAATTGAAATTTAGTCTGCCAGGAGCTGTTGGATCGAAGGTCTTTTC  
 301 GATACCGAGGAATCTGACTTTGCAAGGAATAATTCTTAATCACACCACCCCAACCCCTGA  
 361 ATACACTTCAGGACCCCTGAAACCAACTCGTTGGCTAAATCACAAGAATCTCCCAC  
 10 421 TCATTCCGATTTAGCCAATTAAATATGATATCGGTCTGGAGGCGATAAGGAAATTCT  
 481 ACAAAAAGAGTTATGAATGAGGAAAATAAGGAAAAGAGAGAACTATTTTTAGGTACCC  
 541 TGAAAGAGAACGGAGAAAATTAGAAAAAAACTACTCTCATCTGTACACTGTTCAAA  
 601 GAATATCCnnnnnAATGGTTAGATAATATAAGAAAAGGATAAGTATGATTAAACTGAAAC  
 661 CACGTCGGCAGAACAAAGTGAATTCCCCCTTAAAGAGGAAGTTCGTTCTTAAATATAG  
 15 721 AAAACAAAGAAGTAGTCGCCTCCCTTTAAATGATCTCAGAAAAGAGAAGTAAGTA  
 781 TAAAAGATATTCAAATCTACACAGTCACACTAAATTCTAAATGCTTTCAAT  
 841 TAGCAAATAAAAACAAAAGAAAAAGAAAATTGAAGAAAATCGTTAATAAAACCAT  
 901 TTAAATTCTCAGAAGAAGAGATAAAACAGTTAAAATTGGTCAAACCTTGATTCTTAT  
 961 ACGATGAAGTAAAACAAAAGTTATCTATCTCGGTAAATAAGAAAACCGAAATCTAATA  
 20 1021 ATGATATGCCAAAAGGACAAATCAAATCAAGAAGTTAGACGAAATCGAAAAGAGAT  
 1081 TAAAACAAACTCTGAACGACACAATAATGTGATAGAAGAAACTAAAACACTCAGACTCAT  
 1141 GTTCAGAGTCTCCGATCGTATTGAAAAAATAAAACGTAATAATCAGAGATTCCAGTA  
 1201 AGCGAAATTTCACACTCGCCCCACCTCGATATCATCGAGATGGCGATGGACACCTCA  
 1261 GCATTGATGGAATGGATACTGAGTGATATGATGGATGACAGATGATGAATATAGAAAAC  
 1321 TCACGAAATAACAAATGGCGCTACAGCATATAGAGTAAACATACCGAGGAACAAACAT  
 1381 AAAATTAAATTATCTGGATTACGGGAGTATTAAAAGGCTGGGATAATTACCTCAT  
 1441 GCCAGAACAAAAGAATTATGTTCTAAGCTGTGAAAAATAGAAAACGAAGAAGGAATACC  
 1501 ACTAATGGTGGAAACATTGGTAGCAATAATTCTAACTTATAGGAGATCCAAAGAT  
 1561 TTTTGAGAAAGAACATCTTATTACTTCATAATCTAAGATGTCACCTTAGGTGACTT  
 30 1621 TAGATGGTATTTCAGAAAATTGGTAGCTATGGTTAACAGGGAAAGATTGTAGAGAAC  
 1681 TTTCTGGAAAGAACCGGTTATAGCTGGATTACCGGATATCTTGTGAAAAGGTAAAAGA  
 1741 AAATTACAAAAGGAATGCCAACACCCAATTAAAAGATGTCACCATACGGAAAATAAG  
 1801 TTCAGTTGAAACATCTTATTACTTCATAATCTAAGATGTCACCTTAGGTGACTT  
 1861 AAAAACAGGTGAGAGTCAGGGCATCAAGGAATTAGGGGAATTGGTACTCAATACGGTTA  
 1921 TGAACGAAATACCCCTCCATAAAAAATAAAAGAAAATAGCAAAAGAAGAACAGGGAG  
 1981 AAACAAGCGCTAAAACAAGCGCTAACACAGCACGTTAAACGTTAATT  
 2041 TTAGAAAACCATGAAAGTCTAATGATAAGCCACTATAGCTGTATAATGTGGACGCA  
 2101 TAGGACACATGAAGCGAGACTGTAGACTAAAAGAAAATTAGTAATTGACCATAGTG  
 2161 ATGAATTAAAAGAACAAATGGAAAATTCTGATAATTCTCCAGAAGAGGAAGAAC  
 40 2221 GAAGAATCAATAGGAGATTCTGATTAGAAGTATTGGACATGAGGATAACAATTGTAATT  
 2281 GTGCTCTAAATAAATACGATAAGTAGTGAATTAAAATTGCGTTAGATTGATTGATA  
 2341 AAATTAAATAATCCGGAGAAAAGACCAAGCCTTATAGACATGAAAAGGCTACTCGTTG  
 2401 AAAAGATGAACCCAGTTCATCTCACAAAACCTGAATTATAGGATATGATTAAAG  
 2461 AAATATTGAGAAAAGCGAAAACATCACATAAGAAAATAACCTTAGCGATCTTAAATAGTG  
 45 2521 AAATAAATAAATTAAAAGCGAAATCGAATCTATAAAAGTCGAGCTACAAGAATTAAAG  
 2581 ATAAAATTATACATGAGGAATCCATCTCTGCCAGCAGAAAATTCAACAAGAAGAGGAAG  
 2641 CTAGTAGACCTTCATCAAAGAAATAACATACAAAAGACAAAAGTGGCATGTTAAAG  
 2701 CCCTAGAATTGTTGACCGTTCTGACCGTTCAAGGATGAGTCCTTACCTAA  
 2761 CACAATAAAAACGTTACTCTTAAATATCAAAGGAGAGCTACAAATATCAATGAATGAAT  
 50 2821 GACATTAATATTTCTTAAAGTGTGACCGTTCAAGGATGAGTCCTTACCTAAATATCTGAC  
 2881 TGACTGACATTTCATTTCTGAAATGAGGAGTTATTACGTTAACACCATATAT  
 2941 ATATTTTATCTCAAAGTCAACGAAATTATAAAAAGAATCAATTAAAAAAATTATTCT  
 3001 TTTGCAGAAAAAAATTAAAATATGAAACTCCACACCATATTACCATATTATAAA  
 3061 TATAAAAAACCTCTCACAAATGTGCATTCTGAAATTCTTATGTTGAGAGGATTAATCTC  
 55 3121 TAAAGAAAAAGGTTGAGAAAGGTGCAGCAACA **ATG** TCT CCA TTC TGT AGA  
 1  
 3172 AAC TTT TCA ATG GCA TGG GTG CTT ATG GCA TTT GTG TTG TTT  
 7 N F S M A W V L M A F V L F  
 3214 GCA AAC AGT GCT ATG CCC ACA AAT GGA TCC ACT GTT GGG GTA

	21	A	N	S	A	M	P	T	N	G	S	T	V	G	V			
	3256	AAA	AAC	ATG	TTG	GGT	GGT	AAA	TTG	ATG	CTA	AAC	GTT	TTA	TGT			
	35	K	N	M	L	G	G	K	L	M	L	N	V	L	C			
5	3298	CCC	CAT	ATT	GAT	AAG	CAA	CAC	ATT	ATC	CCG	AAT	GGT	GGT	TCA			
	49	P	H	I	D	K	Q	H	I	I	P	N	G	G	S			
	3340	TTT	GAG	TGG	AAG	TAC	AAT	GGT	GGT	GCT	CCA	CCA	ATA	GGA	CAA			
	63	F	E	W	K	Y	N	G	G	A	P	P	I	G	Q			
10	3382	TCA	CCA	TTC	ATG	TGT	TTC	TTT	CGG	TGG	AAT	AAT	GTT	CAT	CAC			
	77	S	P	F	M	C	F	F	R	W	N	N	V	H	H			
	3424	TCC	CTT	GAT	CTG	TGT	TCA	CCA	AGC	AAG	TAT	ACT	GGT	TGT	GAA			
	91	S	L	D	L	C	S	P	S	K	Y	T	G	C	E			
	3466	AAT	GCC	ATT	TGG	GAA	ATC	AAA	GAA	AAG	CAA	TTT	TGT	AGG	TAC			
	105	N	A	I	W	E	I	K	E	K	Q	F	C	R	Y			
15	3508	AGA	GGT	GGA	CCT	ATT	AAT	TAT	TTT	TGC	TAT	GAC	TGG	GAT	GAT			
	119	R	G	G	P	I	N	Y	F	C	Y	D	W	D	D			
	3550	TAG	TTATATAGATTATT	CATGTT	CATCTCA	ATAAAAAA	ATGACT	TTAGAGT	GATT	CTT								
	3609	AGTTGCTTA	ACATTCTTACAT	ATTCTTA	ACTATTCCGT	CACTACCAC	CCGTA	ACTATAT										
	3669	TTATTTAA	AAATTAGT	TATCTGT	CACAGT	TTTATTTT	TTA	AAAAAGGT	TTATGT	GGATT	AGAAG							
	3729	AGAGATAA	AAATATGT	AGACGGT	CACCAAC	CTTA	ATTTT	GAACT	TATGT	AAGACT	TATATTG							
20	3789	CCAAGA	ATATATGTT	AAACTC	CATTCA	TTAAAG	AGACT	TATATCTC	CATT	TTATG	ATTATGCA							
	3849	AATG	CAATTAG	TTTTTTTC	CATTG	AAAGAATT	CAAAAG	AAAGTT	TATCAT	TTAAAG	TAT							
	3909	CATTA	AAATCACTT	ATATGTT	TTCTTA	ATATC	TT	CTTATTG	TTA	AGAATA	ATTTTT							
	3969	ATCCTT	AAATTAAAG	GTATTACT	ACTTTT	TTTC	CATATCTT	CATT	TTTGAA	ATATT	TTT							
	4029	TAAA	ATT	TCA	ATTTTG	TAACAC	CCC	CAGAAA	AAATAC	ATGTA	ACTAT	CACT	TTTT	TTT	TA			
25	4089	TATT	ACA	AAATT	TGACT	TATAGAA	ATACA	AAATAT	TTAA	AAATATA	AGGTT	CAAA	ACTACA					
	4149	TCCT	AAAGT	CTTT	CAGAC	CCCT	CTGAC	ACATG	TAT	CATCTG	CTG	TAT	TGT	GATA	CAGTC			
	4209	ATCG	CAG	TT	CACA	AGATA	ACAAG	AAA	CCA	AGGGT	AAAGCT	TA	AGA	AAA	AAATTCCATAA			
	4269	CATAT	TTA	ATT	CATG	CAAA	AGAAC	AG	CTA	AGTAA	T	TTT	AAAC	ATT	CTTTAAA			
	4329	TATTG	TTA	AAA	ATT	CAAT	ATCA	TT	CAT	CAT	ATAGACC	ACAC	ATGG	ATCT	TA			
	4389	TTTC	CAAT	CAC	AT	CATTG	GATTT	CATT	TTA	ATC	CTCTG	CTTCC	AGAAG	ACTC	ATT			
	4449	AAGT	ATG	CCCC	TAC	CAAGAG	ACTA	ACAC	CTA	AAAGAG	AAATG	ATCA	AGGTA	AGTT	CAA			
	4509	ACAT	CCA	AAATA	ACGAGT	GCCT	ACAGT	GGGAC	CCAA	ATGT	GTAT	GA	ACTC	CTT	ATCAGCTT			
	4569	CACCAC	CTG	AT	CTT	ATT	CTAT	ATGAC	GTAG	ATC	AGT	GAA	ACTA	GGAGG	ATCTCCGT			
	4629	TAAAC	AT	ATG	TTTT	TAT	ACT	TTA	ATG	CAT	CAA	ACA	ACT	CAC	ATT	CCAA	ATG	
	4689	TATG	AC	CAT	CA	TT	CAT	ACA	TTT	CAT	AT	TAT	ACAT	TATT	GAAT	CACA		
	4749	TAAC	ATT	AAA	ATT	CAT	ACC	ATT	CAAGA	ACT	TTT	CCAAC	AT	CAA	AGCA	AT	TTT	ACTT
	4809	TCAA	ACT	TCAA	AA	AT	ATA	TTT	ATT	TTA	ATA	AA	AGCT					

**FIGURE 5**

142000	TTATCTTATTTCCATATAATTGTTGTTTACTTTCAAAATTAAATTT
141950	TTATATTATCTTTTACAGTTAAAATTAATAAAATGAAACTTTTC
141900	TTAAATGTGTTAAATATAAAATCAAAAAGTTGTTATGGTACATGGC
5 141850	ACAATCTTATAAATTATAATTGAAAACGATACTTATATAATAAAATT
141800	ATCTTAGTTGACATTTTATTAGTGTGTTCAATCATATTGTTGCTT
141750	GATAAGCGTAAAACAAATCAAACCTAACGATACTTATATAATAAAATT
141700	TCTTAGTTGACATTTTATTAGTGTGTTCAATCATATTGTTGCTT
141650	ATAAGCGTAAAACAAATCAAGTAAAGTTGGGCACCTCAATTGTTTAAAAA
10 141600	AAGTTGGGTACCTCAAAATTAATAAGGCTTGTCAAGATTCTTACAAAAA
141550	AAATCTGGAAGAATTATGAAAGAAGGGGGGGAGGGGGGGAGGGGGGG
141500	AAAGTGAAGATGAATATTCAACAAAAGAGGGTAGGCATGATGTTAAGTGAG
141450	TTAAAAAACTATGTTAATGGAGACAATTCTGTAAACAAACCCGTTAAT
141400	TGAAAACGATAGCATTCTCTAACATGAAAACGATATTGTTTATC
15 141350	ATAACTACTCATTAAATTCTGAGTTCAAATCATATAAAAGATTAGGGG
141300	GGTGTATTCAATTAAAGGATTGAAATGATTGTTAAATGACAAATCC
141250	CATGTTATTCAACATGAATTGTAACAAACTTTTAAATCAAGTGTAA
141200	TTAGATTAGTGTGTTAAATGTACAACCAAACCCACTGTTATTGAAAC
141150	ATTTTAAGTAGTGGATTAAATGACTTGAGTGATTGGGTGGGATTGC
20 141100	AGAAAATTCTTAGTTAAGAATTCAAAACATCCAAATCTCATGGTTCAAG
141050	TAGAATTGGGAGAATTAAATAACAAATCTCTAATTACAAAGTC
141000	CCAAATCATTAAAAACTCATAAAAATTAAATGATTCAAATCTCCAG
140950	TTGAATACATCCCTTGGATTAGAGATTGCTCGATTGGGACCTAAG
140900	ATTGAATTGGGATTAGTTAATCGTTACAACAAATGACATCGTAT
25 140850	TATTGTTATAGGAACAATGTCGTTTCAGTTGACATGTATGTTAATAGA
140800	AAATTAACTCTATTAAACGGGATTGCTAACCCATTAAACATGTAACAA
140750	ATGGTCAAGTCATAAAAGTTGGTATTGAAAGTCACGTAAGT
140700	TTGATATTGAAAGTCACATAAAATTGATATCTTATTGTTTC
140650	GACAGACATAAGGATTACATCAATGTTTAATAAAATTAAAGATTATTA
30 140600	TGACATTTCACATTAAATGCCAATGTTTCGAAACCAAGATACTC
140550	AAAATTGACATACCTAACATTCAATCTACATTGTTGACAGCAATTACGT
140500	GCCTTGACACATGGCACATACTGGCAATACATCAATTAAAGGAAAAGG
140450	TAGATTGGATACAATATAATGAAATAAGTGGAAAGGATCATTGACTAC
140400	TTGACTTGTAAACAAACACACAGTATAACTCATCGACATTACAA
35 140350	ACAACATTGTGCTAGCTTAAACTCCCTCTCCTATTCAAAAAA <b>ATG</b>
1	M
140305	GAT ATT CCA AAG CAA TAT CTA TCA CTA TTC ATA TTG
2	D I P K Q Y L S L F I L
140269	ATT ATC TTC ATA ACT ACA AAA TTA TCA CAA GCC GAC
40 14	I I F I T T K L S Q A D
140233	CAT AAA AAC GAC ATT CCA GTT CCC AAC GAT CCA TCA
26	H K N D I P V P N D P S
140197	TCA ACA AAT TCT GTG TTT CCT ACC TCG AAA AGA ACC
38	S T N S V F P T S K R T
140161	GTG GAA ATC AAT AAT GAT CTC GGT AAT CAG CTA ACG
45 50	V E I N N D L G N Q L T
140125	TTA CTG TAT CAT TGT AAA TCA AAA GAC GAT GAT TTA
62	L L Y H C K S K D D D L
140089	GGT AAC CGG ACT CTG CAA CCA GGT GAG TCG TGG TCT
50 74	G N R T L Q P G E S W S
140053	TTT AGT TTC GGG CGT CAA TTC TTT GGA AGG ACG TTG
86	F S F G R Q F F G R T L
140017	TAT TTT TGT AGT TTT AGT TGG CCA AAT GAA TCG CAT
98	Y F C S F S W P N E S H
55 139981	TCG TTC GAT ATA TAT AAA GAC CAT CGA GAT AGC GGC
110	S F D I Y K D H R D S G
139945	GGT GAT AAC AAG TGC GAG AGC GAC AGG TGT GTG TGG
122	G D N K C E S D R C V W
139909	AAG ATA AGA AGA AAC GGA CCT TGT AGG TTT AAC GAT
60 134	K I R R N G P C R F N D

139873	GAA	ACG	AAG	CAG	TTT	GAT	CTT	TGT	TAT	CCT	TGG	AAT	
146	E	T	K	Q	F	D	L	C	Y	P	W	N	
139837	AAA	TCT	TTG	TAT	TGA	CAACAATATGCTGATGTTCTGTCTTTAC							
158	K	S	L	Y	•								
5	139793	GA	CT	ATGGAG	TTT	CATTG	TTT	GAA	ACA	ATA	ATA	AAA	ATT
139743	TCT	ATT	AT	CCA	AG	TT	CC	AA	CT	T	TA	AA	TT
139693	AT	CAT	CTT	AA	GC	CATT	CA	TG	CT	AC	AA	AG	CT
139643	CATT	AAA	AG	CTG	AA	AC	AG	AG	AC	AA	TA	CT	AG
139593	TCT	TT	AT	G	C	A	AC	CA	AT	AC	AC	AA	AG
10	139543	TA	ATA	ATT	C	T	ACT	TT	GG	GT	G	AG	GT
139493	TG	T	G	T	AT	CC	CAT	AC	AC	AT	AT	TT	GT
139443	CAT	G	C	T	CT	CCCC	AC	AT	G	CT	CC	AC	G

卷之三

**FIGURE 6**

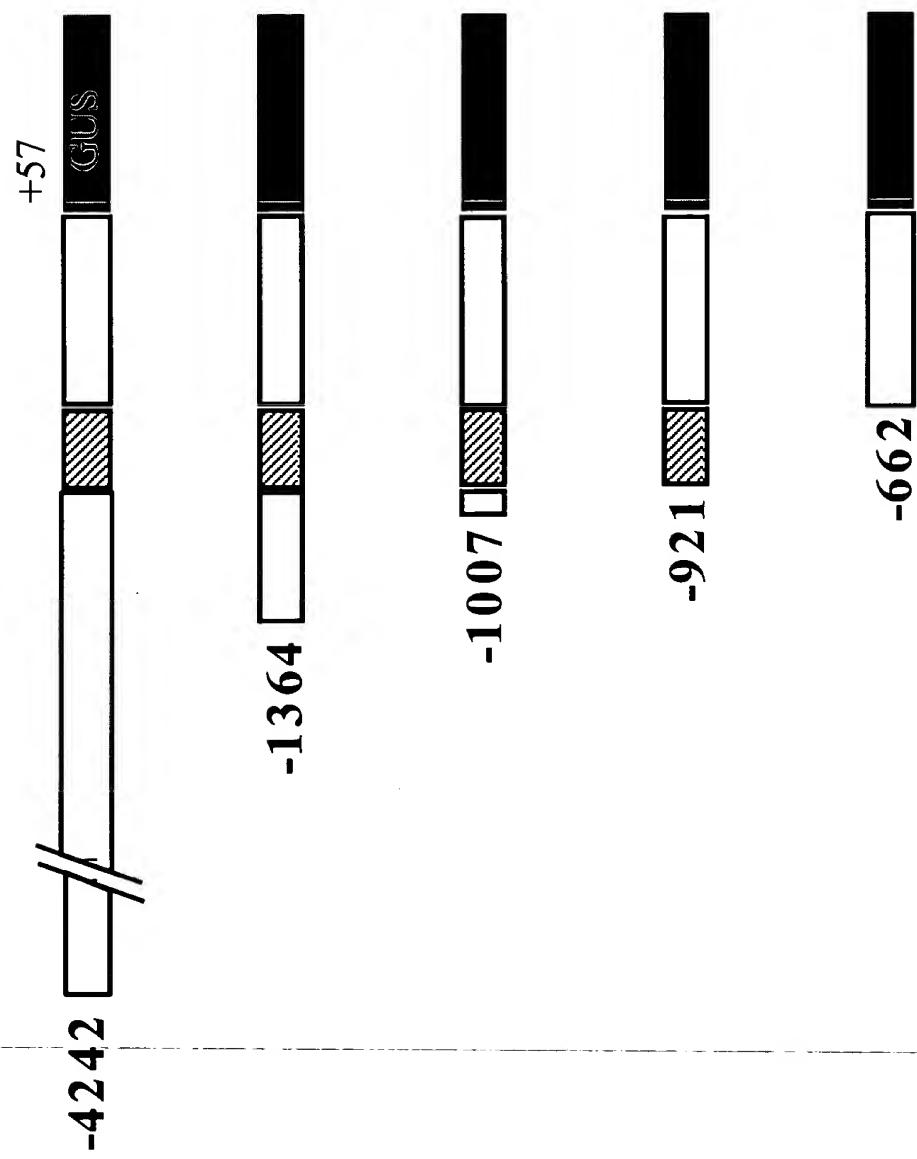
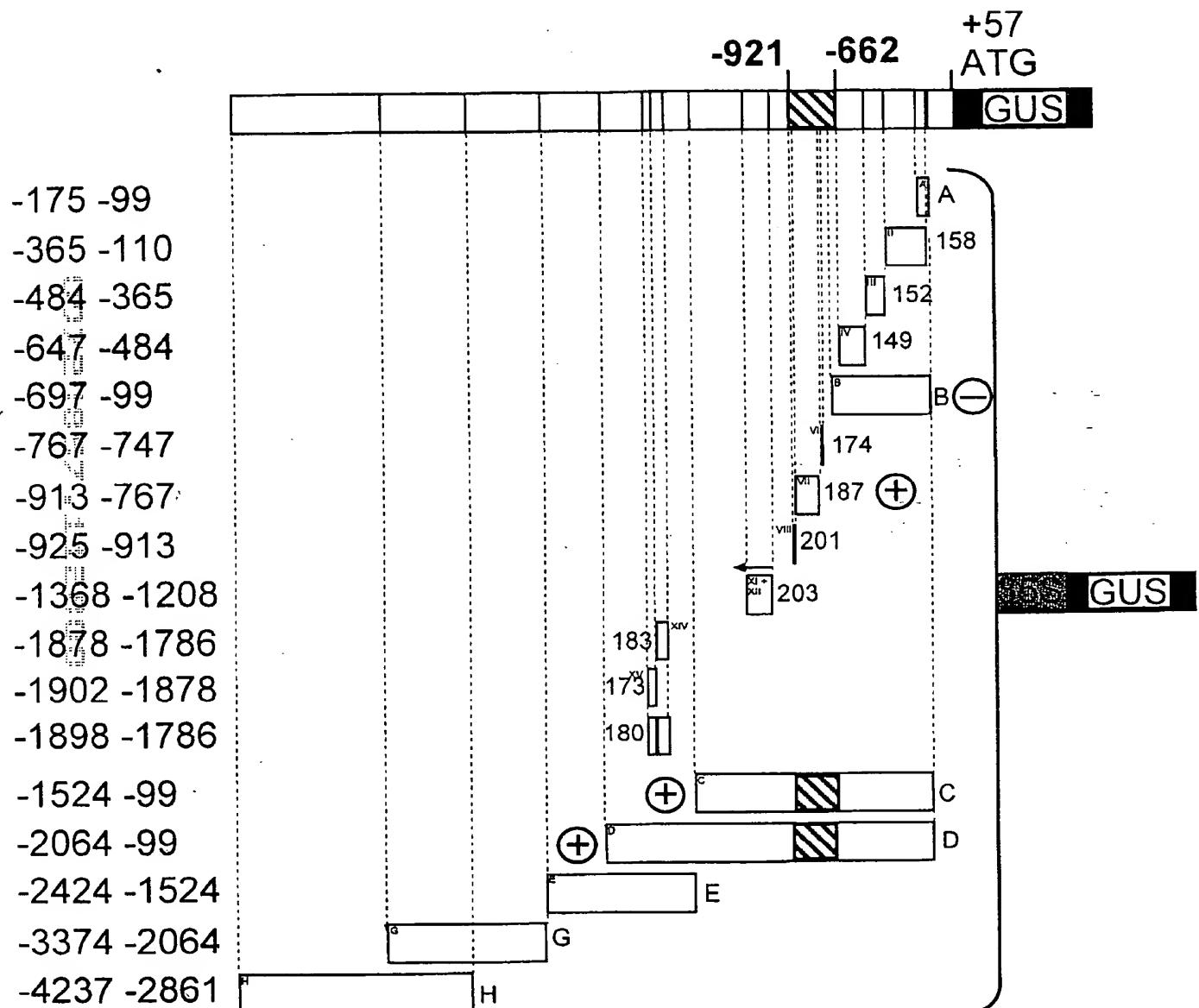


Figure 7

# G564 promoter: Gain of function constructs



⊕ = Gus Activity in Suspensor

⊖ = No Gus Activity in Suspensor

▨ = Region Required for Gus Activity in the Suspensor

## Figure 8

Web Signal Scan Program

Database searched: PLACE  
 URL: <http://www.dna.nffrc.go.jp/hdocs/PLACE/>

This is the sequence you submitted  
 >G564 promoter (-921 to -662), 450 bases, 3D1A0BF4 checksum.  
 TGAAAAGTGAAGAAAACCATGTAATGAAAACAAAATGGCACGACAATCAA  
 AAAAGTTTCACGCAAATTTCTCAAACATTTATAACATTTCATGTT  
 GTGTTTGTCTCAAAGCTAGAAAAACGAAGAGTTACTATTGGTAATGAAA  
 AGCGAAGAAAACCACATAATAAAAACAAAATGGCACGACAATCAAGAAAA  
 AGTTTACACAAAACCTTTTCAAATTTACTATGTTATTCGAAATT  
 TAGAAAAACGAAGAGTTATTAGTAATGAAAAGCGAAGAAAACGTACGT  
 AATAAAAACAAAATGGCACGACAATAAAAAGTTTCACGCAAATTT  
 TCTTGGTGCAGAAAAGTTATATATATAATTAAATTAAATTTCATTACT  
 TTTTCCCTTTTAAAGTTAAATTATTATTTCATTAAAT

Notation: H = A, C, or T  
 R = A or G  
 K = G or T  
 W = A or T

### RESULTS OF YOUR SIGNAL SCAN SEARCH REQUEST

/tmp/signalseqdone.9437: 450 base pairs  
 Signal Database File:

Factor or Site Name	Loc. (Str.)	Signal Sequence	SITE #
-300ELEMENT	site 1 (+)	TGHAAARK	S000122
2SSEEDPROTBANAP	site 101 (-)	CAAACAC	S000143
ACGTABOX	site 296 (+)	TACGTA	S000130
ACGTABOX	site 296 (-)	TACGTA	S000130
AP3SV40	site 159 (-)	TGTGGWWW	S000169
CAATBOX1	site 44 (+)	CAAT	S000028
CAATBOX1	site 189 (+)	CAAT	S000028
CAATBOX1	site 323 (+)	CAAT	S000028
CAATBOX1	site 138 (-)	CAAT	S000028
CANBNNAPA	site 101 (-)	CNAACAC	S000148
CCAATBOX1	site 138 (-)	CCAAT	S000030
CEREGLUBOX2PSLE	site 55 (-)	TGAAAAC	S000033
CEREGLUBOX2PSLE	site 201 (-)	TGAAAAC	S000033
CEREGLUBOX2PSLE	site 333 (-)	TGAAAAC	S000033
DOFCOREZM	site 4 (+)	AAAG	S000265
DOFCOREZM	site 53 (+)	AAAG	S000265
DOFCOREZM	site 112 (+)	AAAG	S000265
DOFCOREZM	site 149 (+)	AAAG	S000265
DOFCOREZM	site 199 (+)	AAAG	S000265
DOFCOREZM	site 282 (+)	AAAG	S000265
DOFCOREZM	site 331 (+)	AAAG	S000265
DOFCOREZM	site 364 (+)	AAAG	S000265
DOFCOREZM	site 419 (+)	AAAG	S000265
DOFCOREZM	site 216 (-)	AAAG	S000265
DOFCOREZM	site 399 (-)	AAAG	S000265
DOFCOREZM	site 408 (-)	AAAG	S000265
GT1CONSENSUS	site 120 (+)	GRWA	S000198
GT1CONSENSUS	site 141 (+)	GRWA	S000198
GT1CONSENSUS	site 196 (+)	GRWA	S000198
GT1CONSENSUS	site 253 (+)	GRWA	S000198
GT1CONSENSUS	site 69 (-)	GRWA	S000198
GT1CONSENSUS	site 90 (-)	GRWA	S000198
GT1CONSENSUS	site 347 (-)	GRWA	S000198
GT1CONSENSUS	site 388 (-)	GRWA	S000198
GT1CONSENSUS	site 436 (-)	GRWA	S000198

GT1CONSENSUS	site	218	(-)	GRWAAW	S000198
GT1CONSENSUS	site	401	(-)	GRWAAW	S000198
GT1CONSENSUS	site	402	(-)	GRWAAW	S000198
MAMMALENHAN	site	158	(-)	GTGGTTTK	S000121
MARTBOX	site	324	(-)	TTWTWTWT	S000067
MRE1	site	356	(-)	TGCRCNC	S000068
NTBBF1ARROLB	site	418	(-)	ACTTTA	S000273
POLASIG1	site	168	(+)	AATAAA	S000080
POLASIG1	site	301	(+)	AATAAA	S000080
POLASIG1	site	324	(+)	AATAAA	S000080
POLASIG1	site	237	(-)	AATAAA	S000080
POLASIG1	site	411	(-)	AATAAA	S000080
POLASIG3	site	268	(-)	AATAAT	S000088
POLASIG3	site	427	(-)	AATAAT	S000088
POLASIG3	site	430	(-)	AATAAT	S000088
POLASIG3	site	433	(-)	AATAAT	S000088
POLLEN1LELAT52	site	11	(+)	AGAAA	S000245
POLLEN1LELAT52	site	119	(+)	AGAAA	S000245
POLLEN1LELAT52	site	156	(+)	AGAAA	S000245
POLLEN1LELAT52	site	195	(+)	AGAAA	S000245
POLLEN1LELAT52	site	252	(+)	AGAAA	S000245
POLLEN1LELAT52	site	289	(+)	AGAAA	S000245
POLLEN1LELAT52	site	362	(+)	AGAAA	S000245
POLLEN1LELAT52	site	71	(-)	AGAAA	S000245
POLLEN1LELAT52	site	349	(-)	AGAAA	S000245
PYRIMIDINEBOXHV	site	400	(+)	TTTTTTCC	S000298
RAV1AAT	site	97	(-)	CAACA	S000314
ROOTMOTIFTAPOX1	site	374	(+)	ATATT	S000098
SEF4MOTIFGM7S	site	170	(-)	RTTTTTR	S000103
SP8BFIBSP8BIB	site	134	(+)	TACTATT	S000184
TATABOX2	site	81	(-)	TATAAAT	S000109
TATABOX3	site	375	(+)	TATTAAT	S000110
TATABOX4	site	368	(-)	TATATAA	S000111
TATABOX5	site	238	(+)	TTATTT	S000203
TATABOX5	site	412	(+)	TTATTT	S000203
TATABOX5	site	434	(+)	TTATTT	S000203

-----

For more information about the SignalScan Program, please contact Dr Dan S. Prestridge Tele: (612) 625-3744 Advanced Biosciences Computing Center. E-mail: danp@biosci.umn.edu 1479 Gortner Ave. University of Minnesota St. Paul, MN 55108 The TFD data is at the gopher site, gopher://genome-gopher.stanford.edu. For more information about the WebSignalScan service, please contact Meena Sakharkar, meena@biomed.nus.sg, BioInformatics centre, NUS .

Database Searched: PlantCARE  
URL: <http://sphinx.rug.ac.be:8080/PlantCARE/>

Sequence submitted:

>G564 promoter (-921 to -662) 11/21/00

+ GAAAAAGTCAA GAAAACCATG TAATGAAAAC AAAATGGCAC GACAATCAA AAAAGTTTC ACGAAAATT  
+ TTCTTCAAAA TTTATAACAT TTTCATGTTG TGTTTGTTC AAAGCCTAGA AAAACGAAGA GTTACTATTG  
+ GTAATGAAAA GCGAAGAAAA CCACATATA AAAACAAAAT GGCACGACAA TCAAGAAAAA GTTTCACAC  
+ AAAACTTTT TCAAATTTA CTATGTTAT TTCGAAATTT AGAAAAACGA AGAGTTATTA TTAGTAATGA  
+ AAAGCGAAGA AAACTACGTA ATAAAAAACAA ATGGCACG ACAATAAAA AAGTTTCAC GCAAATTTT  
+ CTTGGTGCAG AGAAAGTTAT ATATATTAAT TAATTAATTT TCATTTACTT TTTCCCTTT TTATTTAAA  
+ GTTAAATTAT TATTATTTTC ATTTAAAAA

9 - CTTTCACTT CTTTGGTAC ATTACTTTG TTTTACCGTG CTGTTAGTTT TTTCAAAAG TGCCTTTAA  
10 - AAGAAGTTT AAATATTGTA AAAGTACAAC ACAAAACAAAG TTTCGGATCT TTTGCTTCT CAATGATAAC  
11 - CATTACTTT CGCTTCTTT GGTGTATTAT TTTGTTTA CCGTGCTGTT AGTTCTTTT CAAAAGTGTG  
12 - TTTGAAAAA AGTTTAAAT GATACAAATA AAGCTTAAAC TCTTTTGCT TCTCAATAAT AATCATTACT  
13 - TTTCGCTTCT TTTGATGCAT TATTTTTGT TTTACCGTGC TGTTATTTT TTCAAAAGTG CGTTTAAAAA  
14 - GAACCACGGG TCTTCATAATA TATATAATTA ATTAATTAAA AGTAAATGAA AAAAGGGAAA AATAAAATTT  
15 - CAATTAAATA ATAATAAAAG TAAATTTT

16 3-AF1\_binding\_sit

Site Name	Organism	Position	Strand	Core	Matrix	sequence
				simil.	simil	
3-AF1_binding_sit	ST	260	+	1.000	0.860	AAGAGttatt

Function:

AAGAA-motif

Site Name	Organism	Position	Strand	Core	simil.	Matrix	simil	sequence
AAGAA-motif	Avena sativa	6	+	1.000	0.903	gtgAAGAA		
AAGAA-motif	Avena sativa	151	+	1.000	0.870	gcaAAGAA		
AAGAA-motif	Avena sativa	284	+	1.000	0.870	gcaAAGAA		

Function:

ABRE

Site Name	Organism	Position	Strand	Core	simil.	Matrix	simil	sequence

ABRE	Hordeum vulgare	293	+	1.000	0.854	actACGTaat		
------	-----------------	-----	---	-------	-------	------------	--	--

Function: cis-acting element involved in the abscisic acid responsiveness

ACE

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
ACE	<i>Petroselinum crispum</i>	293	+	1.000	0.908	actACGTaat

Function: cis-acting element involved in light responsiveness

• AE-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
AE-box	<i>Arabidopsis thaliana</i>	67	-	1.000	0.852	AGAAaatt
AE-box	<i>Arabidopsis thaliana</i>	345	-	1.000	0.852	AGAAaatt
AE-box	<i>Arabidopsis thaliana</i>	361	+	1.000	0.852	AGAAagtt

Function: part of a module for light response

AT1-motif

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
AT1-motif	<i>Solanum tuberosum</i>	409	+	1.000	0.859	ttttATTTtaaa

Function: part of a light responsive module

Box\_4

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
Box_4	PC	375	+	1.000	1.000	ATTAat
Box_4	PC	379	+	1.000	1.000	ATTAat
Box_4	PC	383	-	1.000	1.000	ATTAat

Function:

Box\_I

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
Box_I	PS	107	+	1.000	1.000	TTTCaaa
Box_I	PS	203	+	1.000	0.857	TTTCaca
Box_I	PS	219	+	1.000	1.000	TTTCaaa
Box_I	PS	240	+	1.000	0.857	TTTCgaa
Box_I	PS	241	-	1.000	0.857	TTTCgaa
Box_I	PS	249	-	1.000	0.857	TTTCtaa

Function:

Box\_II

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
-----------	----------	----------	--------	-------------	--------------	----------

Box_II	ST	139	+	1.000	0.889	TGGTaatga
Box_II	AT	161	+	1.000	0.954	CCACataat

Function:

CAAT-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
CAAT-box	Hordeum vulgare	43	+	1.000	1.000	CAAT
CAAT-box	Arabidopsis thaliana	137	-	1.000	1.000	aCCAAt
CAAT-box	Hordeum vulgare	188	+	1.000	1.000	CAAT
CAAT-box	Hordeum vulgare	322	+	1.000	1.000	CAAT
CAAT-box	Arabidopsis thaliana	351	-	1.000	0.857	aCCAAg

Function: common cis-acting element in promoter and enhancer regions

ERE

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
ERE	Dianthus caryophyllus	239	+	1.000	0.875	ATTTcgaa
ERE	Dianthus caryophyllus	241	-	1.000	0.875	ATTTcgaa
ERE	Dianthus caryophyllus	413	+	1.000	0.875	ATTTt.aaa
ERE	Dianthus caryophyllus	441	+	1.000	0.875	ATTTaaaa
ERE	Dianthus caryophyllus	442	-	1.000	0.875	ATTTt.aaa

Function: ethylene-responsive element

G-box

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
G-box	Zea mays	17	+	0.842	0.870	CATGta
G-box	Zea mays	38	+	1.000	0.903	CACGac
G-box	Zea mays	94	+	0.842	0.886	CATGtt
G-box	Zea mays	183	+	1.000	0.903	CACGac
G-box	Zea mays	317	+	1.000	0.903	CACGac

Function: cis-acting regulatory element involved in light responsiveness

GC-repeat

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
GC-repeat	Oryza sativa	351	-	1.000	1.000	gCACCAag

Function: ?

HSE

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
HSE	Brassica oleracea	49	+	0.944	0.878	aAAAAagtt
HSE	Brassica oleracea	50	+	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	52	-	0.944	0.874	gAAAAActtt
HSE	Brassica oleracea	66	-	1.000	0.978	aGAAAattt
HSE	Brassica oleracea	77	-	0.833	0.868	aTAAAAtttt
HSE	Brassica oleracea	87	-	1.000	0.853	tGAAAatgl
HSE	Brassica oleracea	196	+	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	198	-	0.944	0.874	gAAAActtt
HSE	Brassica oleracea	210	+	0.944	0.874	cAAAActtt
HSE	Brassica oleracea	212	-	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	213	-	0.944	0.878	aAAAAagtt
HSE	Brassica oleracea	327	+	0.944	0.878	aAAAAagtt
HSE	Brassica oleracea	328	+	0.944	0.912	aAAAAgttt
HSE	Brassica oleracea	330	-	0.944	0.874	gAAAActtt
HSE	Brassica oleracea	344	-	1.000	0.978	aGAAAattt
HSE	Brassica oleracea	361	+	1.000	0.888	aGAAAgtta
HSE	Brassica oleracea	385	-	1.000	0.853	tGAAAatta

Function: cis-acting element involved in heat stress responsiveness

I-box

Site	Name	Organism	Position	Strand	Core	Matrix	sequence
I-box		Pisum sativum	93	-	0.857	0.883	aACATga
I-box		Pisum sativum	162	+	0.857	0.883	cACATaa
I-box		Solanum tuberosum	163	-	1.000	1.000	tATTAtgt
I-box		Pisum sativum	237	-	0.857	0.941	gAAATaa
I-box		Pisum sativum	367	-	1.000	1.000	tATATaa
I-box		Pisum sativum	372	+	1.000	0.941	tATATta
I-box		Pisum sativum	391	-	0.857	0.941	tAAATga
I-box		Pisum sativum	411	-	0.857	0.883	aAAATaa
I-box		Pisum sativum	423	+	0.857	0.883	tAAATta
I-box		Solanum tuberosum	424	-	1.000	0.903	aATAAAttt
I-box		Arabidopsis thaliana	426	-	1.000	0.863	aATAAAtaat
I-box		Arabidopsis thaliana	429	-	1.000	0.863	aATAAAtaat
I-box		Solanum tuberosum	431	+	1.000	0.951	tATTAttt
I-box		Pisum sativum	433	-	0.857	0.883	aAAATaa
I-box		Pisum sativum	439	-	0.857	0.941	tAAATga
Function: part of a light responsive element							

#### P-box

Site	Name	Organism	Position	Strand	Core	simil.	Matrix	simil.	sequence
P-box		Oryza sativa	406	+	1.000	0.857			CCTTttt
Function: gibberellin-responsive element									

#### Prolamin\_box

Site	Name	Organism	Position	Strand	Core	simil.	Matrix	simil.	sequence
Prolamin-box	Oryza sativa		145	+	1.000	0.913			tgaAAAGC
Prolamin-box	Oryza sativa		278	+	1.000	0.913			tgaAAAGC
Function: cis-acting regulatory element associated with GCN4									

#### TATA-box

Site	Name	Organism	Position	Strand	Core	simil.	Matrix	simil.	sequence
TATA-box	Daucus carota		79	-	1.000	1.000			TATAaatt
TATA-box	Brassica juncea		80	-	1.000	1.000			TATAaat
TATA-box	Helianthus annuus		81	-	1.000	1.000			TATAaaa
TATA-box	Brassica oleracea		82	+	1.000	0.908			tTATAaac
TATA-box	Brassica napus		83	-	1.000	0.892			gtTATA
TATA-box	Oryza sativa		117	+	0.818	0.912			TAGAaaa
TATA-box	Oryza sativa		169	+	0.818	0.872			TAAAaac
TATA-box	Zea mays		248	+	0.909	0.879			TTTAgaaa
TATA-box	Oryza sativa		250	+	0.818	0.912			TAGAaaaa
TATA-box	Oryza sativa		302	+	0.818	0.912			TAAAaaa
TATA-box	Oryza sativa		325	+	0.818	0.912			TAAAaaa
TATA-box	Daucus carota		364	-	1.000	0.863			TAAAaaa
TATA-box	Brassica juncea		365	-	1.000	0.857			TATAaact
TATA-box	Zea mays		366	-	1.000	0.879			TATATAaac

TATA-box	Oryza sativa	367	-	1.000	0.956	TATAtaa
TATA-box	Oryza sativa	368	+	1.000	0.929	TATAtat
TATA-box	Oryza sativa	369	-	1.000	0.929	TATAtat
TATA-box	Solanum tuberosum	370	-	1.000	1.000	TATAta
TATA-box	Glycine max	372	+	1.000	0.891	TATAtt
TATA-box	Oryza sativa	407	-	0.818	0.872	TAAAaag
TATA-box	Zea mays	413	-	0.909	0.879	TTTAAAat
TATA-box	Zea mays	442	+	0.909	0.879	TTTAAAat

Function: core promoter element around -30 of transcription start

• TC-rich\_repeats

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
TC-rich_repeats	NT	7	-	1.000	0.952	gTTTTcttca
TC-rich_repeats	NT	68	+	1.000	1.000	aTTTTcttca
TC-rich_repeats	NT	152	-	1.000	0.909	gTTTTcttcg
TC-rich_repeats	NT	191	-	1.000	0.885	tTTTTcttga
TC-rich_repeats	NT	248	-	1.000	0.914	tTTTTctaaa
TC-rich_repeats	NT	285	-	1.000	0.909	gTTTTcttcg
TC-rich_repeats	NT	346	+	1.000	0.915	aTTTTcttgg

WUN-motif

Site Name	Organism	Position	Strand	Core simil.	Matrix simil	sequence
WUN-motif	Brassica oleracea	18	-	1.000	0.948	tCATTacat
WUN-motif	Brassica oleracea	139	-	1.000	1.000	tCATTacca
WUN-motif	Brassica oleracea	237	+	0.857	0.948	tTATTtcga
WUN-motif	Brassica oleracea	242	-	1.000	1.000	aAATTtcga
WUN-motif	Brassica oleracea	272	-	1.000	0.948	tCATTacta
WUN-motif	Brassica oleracea	296	-	0.857	0.948	tTATTacgt

Function: wound-responsive element